



Specialization and Variety in Repetitive Tasks: Evidence from a Japanese Bank

Citation

Staats, B., and F. Gino. "Specialization and Variety in Repetitive Tasks: Evidence from a Japanese Bank." *Management Science* 58, no. 6 (2012): 1141–1159.

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:10996799>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Specialization and Variety in Repetitive Tasks: Evidence from a Japanese Bank

Bradley R. Staats
University of North Carolina at Chapel Hill
Campus Box 3490, McColl Building
Chapel Hill, NC 27599-3490
Tel: 919.962.7343
Fax: 919.962.6949
bstaats@unc.edu

Francesca Gino
Harvard Business School
Harvard University, Baker Library
Boston, MA 02163
Tel: 617.495.0875
Fax: 617.496.4191
fgino@hbs.edu

Forthcoming, *Management Science*

Acknowledgments

We are grateful to Masamoto Yashiro, Jay Dvivedi, Michiyuki Okano, Pieter Franken, Yuki Kimura, Aiko Suga, and numerous other individuals at Shinsei Bank for their investments of time and attention to this project, without which the work would not have been possible. We thank Christian Terwiesch, the associate editor, and the reviewers for substantive comments that significantly shaped this manuscript. Ethan Bernstein, David Brunner, Jonathan Clark, Adam Grant, Dave Hofmann, Rob Huckman, Diwas KC, Saravanan Kesavan, Ann Maruchek, Lamar Pierce, Gary Pisano, Bill Simpson, Jay Swaminathan, Harvey Wagner, and participants at the Conference on Behavioral Research in Operations Management provided helpful comments on earlier drafts of this paper. We also gratefully acknowledge support from the University of North Carolina's Center for International Business Education and Research. All errors are our own.

Abstract

Sustaining operational productivity in the completion of repetitive tasks is critical to many organizations' success. Yet research points to two different work-design-related strategies for accomplishing this goal: *specialization* to capture the benefits of repetition and *variety* (i.e., working on different tasks) to keep workers motivated and provide them opportunities to learn. In this paper, we investigate how these two strategies may bring different productivity benefits over time. For our empirical analyses, we use two and a half years of transaction data from a Japanese bank's home loan application-processing line. We find that over the course of a single day, specialization, as compared to variety, is related to improved worker productivity. However, when we examine workers' experience across a number of days, we find that variety helps improve worker productivity. Additionally, we show that part of this benefit results from workers' cumulative experience with changeovers. Our results highlight the need for organizations to transform specialization and variety into mutually reinforcing strategies rather than treating them as mutually exclusive. Overall, our paper identifies new ways to improve operational performance through the effective allocation of work.

Key Words: Learning, Motivation, Productivity, Specialization, Variety, Work Fragmentation

1. Introduction

From Adam Smith's (1776) pin factory and Frederick Taylor's (1911) coal yards to present-day factories in China, call centers in India, and fast food restaurants and banks in the United States (Upton and Margolis 1992; Huckman and MacCormack 2009), sustaining operational productivity in the completion of repetitive tasks is key to many organizations' success. A long line of work in empirical operations has examined this question, exploring factors that affect productivity (Fisher and Ittner 1999; KC and Terwiesch 2009), learning (Lapr   and Tsikriktsis 2006; Boh, Slaughter and Espinosa 2007) and quality (Lapr  , Mukherjee and Wassenhove 2000; Huckman and Pisano 2006). One tool that managers have to affect productivity is task allocation; however, the most appropriate approach for allocating tasks is unclear, as scholars have long debated whether productivity is higher when work is specialized or varied. On one hand, researchers argue for the productivity benefits of *specialization*: namely, helping workers gain skill in a particular task (Newell and Rosenbloom 1981; Argote 1999) and reducing changeovers (Cellier and Eyrolle 1992; Schultz, McClain and Thomas 2003). On the other hand, scholars suggest that executing a *variety* of different tasks improves performance since workers' motivation and engagement with the job increases (Herzberg 1968; Hackman and Oldham 1976) and the knowledge they gain can be applied to other tasks (Schilling et al. 2003; Wiersma 2007; Narayanan, Balasubramanian and Swaminathan 2009).

In this paper, we seek to understand how work can be structured effectively across tasks and over time in order to improve operational performance. We posit that, in order to disentangle the effects of specialization and variety, it is necessary to consider the distinct benefits that each approach provides *over time*. In other words, it is possible to evaluate the operational implications of a specialized or varied task

assignment strategy over a short time period (such as one day) or a longer time period (e.g., many days), and the mechanisms through which each strategy affects performance suggest differential benefits.

When a worker executes many tasks during a short time period, specialization helps to complete the current task quickly (Newell and Rosenbloom 1981; Argote 1999) and limits costly changeovers (Cellier and Eyrolle 1992; Schultz et al. 2003). At the same time, over a short time period, variety can be distracting and, as a result, mixing the two strategies may negatively impact workers' current productivity (Allport, Styles and Hsieh 1994; Monsell 2003). We propose, however, that this tradeoff between specialization and variety involves different costs and benefits for productivity over longer time periods. By completing different, though related, task types, a worker may identify new best practices over time and transfer those practices from one task to another (Schmidt 1975; Ichniowski and Shaw 1999; Tucker, Nembhard and Edmondson 2007). Additionally, the motivational benefits of variety (Hackman and Oldham 1976; Fried and Ferris 1987) are more likely to be salient when a worker has completed a task a number of times (Ortega 2001). Although prior work examining the individual-level productivity effects of specialization and variety over the longer term has investigated the possible direct effects of variety (Boh et al. 2007; Narayanan et al. 2009), none has considered complementarities that variety may offer over time (Lindbeck and Snower 2000). In other words, we are aware of no prior research that has empirically examined whether the returns to specialization are *increasing* in the amount of varied experience over the long-term. Building on prior work on the trade-off between specialization and variety, this paper addresses the following research questions: (1) How do these two strategies affect productivity in the *short-term* and is there a combined effect between them? (2) How do specialized and varied experience affect productivity in the *long-term* and is there a combined effect between the two?

We also examine one particular mechanism by which variety may aid worker performance over time: learning from experience with changeovers. Prior work notes that changing tasks results in costly cognitive setups as workers acclimate to the new activity (Cellier and Eyrolle 1992; Allport et al. 1994; Schultz et al. 2003). However, as a worker completes more varied tasks, she gains experience in setups that may help her to change tasks more effectively. Thus, our final research question is: Do workers exhibit learning in setups and changeovers?

We address these questions by examining workers' productivity during the completion of repetitive, procedural tasks. Procedural tasks "involve series of discrete motor responses (responses with a distinct beginning and end). The responses themselves are easy to execute; it is deciding what responses to make and in what sequence that pose the main problems for the learner" (Schendel and Hagman 1982: 605). When completing procedural tasks, a worker must exert herself both physically and mentally. Examples of such tasks include common operational processes: manufacturing assembly line operations, hospital transporters, and data entry (the context for this paper). While procedural tasks encompass many

types of work that characterize modern organizations, they do not include “knowledge work” (e.g., scientists or surgeons, Drucker 1999).

Performance in executing procedural tasks depends on the combination of the effort exerted as well as a worker’s learning-by-doing. Many procedural tasks are self-paced and characteristics of either the work context or the worker may affect the effort exerted (Staw 1980; KC and Terwiesch 2009). Additionally, problem-solving by front-line workers completing procedural tasks can yield significant operational improvements (e.g., Spear and Bowen 1999); however, workers’ problem solving efforts are limited by the scope of their task design.

In this paper, we examine worker productivity on procedural tasks by using two and a half years of transaction data from the home loan application processing line at Shinsei Bank, a mid-sized Japanese bank. The bank’s data entry operations offer an ideal setting for several reasons. First, they are highly repetitive – tasks take only a few minutes to complete, and an average worker completes over a hundred tasks per day. Second, tasks are procedural as a worker exerts physical effort (e.g., typing) and mental effort (e.g., devising faster ways to complete the work) to execute the necessary steps. Third, since tasks are self-paced, a worker’s performance is determined by both the effort she puts forth and the learning-by-doing that occurs. Finally, as a result of deploying a new information technology system (Upton and Staats 2008), task-level performance was tracked for each of the 111 workers on the line.

In the next section, we motivate our hypotheses. We then present our data and empirical results before discussing implications of the findings and offering concluding remarks.

2. Specialization and Variety

The concept of specialization has played a central role in developing the field of operations management. The Industrial Revolution led to large-scale operations, creating the need to identify ways to simplify these often complex processes (Skinner 1985). Frederick Taylor stepped into this gap with his principles of Scientific Management, which involved breaking down a task, optimizing the constituent steps, and then focusing workers on repeatedly executing the task (Taylor 1911).¹ Specialization benefits individual workers’ productivity since work on the same task over time imparts knowledge related to the task that is likely to improve a worker’s performance (Newell and Rosenbloom 1981; Argote 1999; Huckman and Pisano 2006). The knowledge gained may cover many topics, including the specific set of steps to follow, the specialized tools used, and the customer being served (Lapr  and Nembhard 2010).

¹ While Taylor’s work concentrated on benefits of specialization at the individual level, subsequent work in operations management has examined this topic at the operating unit level, referring to the topic as “focus” (Skinner 1974). This work generally supports the value of focus (Lapr  and Tsikriktsis 2006; Tsikriktsis 2007; Huckman and Zinner 2008), but does not always do so (MacDuffie, Sethuraman and Fisher 1996; Mukherjee, Mitchell and Talbot 2000). Recent work unpacks focus further, examining the impact of related activities on focus (Clark and Huckman 2011) as well as the different possible types of focus and their effect on performance (KC and Terwiesch 2011).

Specialization not only creates conditions that may foster learning, it also avoids costs that may arise from varied experience.² A large body of research in the operations management literature has examined scheduling and inventory decisions using analytical tools to minimize costly setups and changeovers (e.g., Bahl, Ritzman and Gupta 1987; Allahverdi, Gupta and Aldowaisan 1999). Further work tackled the problem by considering how to decrease the time required for setups to eliminate waste (e.g., Shingo 1989; Tzur 1996). More recently, studies have proposed that not only do machines require setups and changeovers, but so, too, do people (Simons and Russell 2002; Schultz et al. 2003).

While these learning benefits of specialization and costs of variety related to setups point toward an overall benefit of specialization, work in organizational behavior has identified potential costs to this approach, as well. Much of this research traces its roots to fieldwork that documented the cognitive toll on workers who repeatedly execute the same tasks over time (Roethlisberger and Dickson 1934; Roy 1959). While repeated experience offers opportunity for learning, it also introduces the possibility of challenges with motivation and boredom (Fisher 1993). When a task is repetitive, familiar, or dull, workers are more likely to experience only low levels of cognitive arousal and, as a result, might disengage from the task (Warr 2007). Alternatively, they could engage in behaviors that, while raising their arousal levels, also detract from job effectiveness (Vroom 1964; Scott 1966; Hackman 1969). Thus, with repetition of the same task, not only might workers be less likely to identify new ways to improve performance, but they also may lose motivation, resulting in decreased performance.

For these reasons, organizational behavior research on job design and motivation stresses the need for task variety (Hackman and Oldham 1976; Ichniowski and Shaw 1999; Humphrey, Nahrgang and Morgeson 2007). Changing the task may increase workers' mental stimulation or arousal, as well as their task engagement, thus improving performance (Langer 1989). Additionally, task variety can create the opportunity for knowledge transfer between tasks, which may result in learning (Schilling et al. 2003; Narayanan et al. 2009). For example, a worker may recognize that a step used in completing Task A may improve her productivity in completing Task B. Additionally, by completing Task A and Task B, a worker may recognize a higher order principle that affects both tasks.

Given the tension between specialization and variety, the question remains: how should repetitive, procedural tasks be assigned to workers? We propose that temporal considerations affect the necessary balance between these two strategies.

2.1 Specialization and Variety in the Short Term

We first consider worker task assignment over a short time period. Given our setting, we use one

² This paper examines this possibility at the individual level; however, work in operations management at the level of a plant considers the costs of variety that arise from the added operational complexity and challenges of assigning workers, given the variability in task-completion time (e.g., MacDuffie et al. 1996; Fisher and Ittner 1999).

day to capture a “short” time period. In fact, within our context, a worker completes many tasks that typically last a few minutes and do not carry over from one day to the next. We note that short- and long-term time periods would differ in other operational contexts. For instance, in a context for which tasks generally last only several seconds, a “short” time period might be comparatively smaller (e.g., an hour), while a setting with more tasks at least five hours long would have a comparatively longer “short” time period (e.g., perhaps a week). Since our research study uses one day to capture the “short” term, anything over one day captures the “long” term.

Over a short time period, we expect specialization to dominate variety for several reasons. First, specialization over the course of a single day allows a worker to limit changeovers. When a worker switches tasks, she needs to relearn or at least reacquaint herself with key relevant processes (Bailey 1989; Nembhard and Osothsilp 2001). This reacquainting effect is analogous to setting up a machine, likewise resulting in decreased productivity. Consistent with this argument, laboratory studies find that switching tasks does indeed worsen individual performance (Cellier and Eyrolle 1992; Allport et al. 1994; Schultz et al. 2003).

Second, specialization over the course of a day offers potential learning benefits, as well. With repetition, a worker not only gains mastery of the individual steps in a task, but also may see how the pieces fit together and recognize opportunities for improvement (Bohn 2005; Bohn and Lapré 2011). For example, in the context of data entry, a worker might recognize that the current task requires more frequent shifting of her field of vision from the data-entry form to the computer screen than did a prior task (e.g., to input information to only one field at a time as opposed to two fields at a time, due to the nature of data being entered).

Third, benefits of specialization may hold even if a worker executed a similar task many days before. Not only is there a risk that she may have forgotten prior knowledge (Bailey 1989; Argote, Beckman and Epple 1990), but the same day experience should help her move all relevant knowledge into short-term or working memory for easier access (Baddeley 1992). This could function in a manner similar to a computer’s moving programs or data from long-term memory into a more rapidly accessible cache to improve performance. Same-day experience offers workers not only mental benefit, but physical benefit, as well: as individuals begin to execute tasks, they gain muscle memory to improve productivity. Altogether, by executing the same task repeatedly over a short period, a worker may get into a groove, steadily improving performance (Quinn 2005).

Despite these benefits of specialization, organizational behavior research on motivation and job design suggests that task variety is necessary to maintain worker productivity (Fried and Ferris 1987; Humphrey et al. 2007). However, we hypothesize that these gains in motivation and learning are likely to be dominated by the costs of task variation within a single day. First, while changing tasks may provide

some motivational benefit, we posit that any productivity improvement within a single day is likely to be small and overshadowed by task-switching costs. Within a single day, in fact, individuals may not have enough time to get acquainted with a task before switching to a different one, thus incurring cognitive costs (Allport et al. 1994; Rogers and Monsell 1995). Second, although variety can lead to learning, we expect that such learning is unlikely to manifest itself in a substantial way during a single day. Recognizing opportunities for performance improvement typically requires reflection (Argyris and Schön 1978), a process that is difficult to do during the course of a busy day filled with repetitive tasks. Thus, while variety during the course of a day may still improve performance on a focal task, it is unlikely to do so more than would a specialized strategy. Thus, we hypothesize that:

HYPOTHESIS 1: *Same-day, same-task experience has a greater positive effect on worker productivity than does same-day, different-task experience.*

Next, we consider the combined effects of specialization and variety within the course of a single day. Psychological research suggests not only that changing tasks impairs performance due to the new task's initial setup cost, but also that a longer-term residual cost to variety may exist (Allport et al. 1994; Monsell 2003). In our context, this suggests that the interaction of varied and specialized experience may have a negative impact on performance, due to the potential distracting effects of variety.

The reasons for this have to do with how the brain processes task change. When an individual switches between tasks she must move the necessary steps and processes for completing the new task into her working memory (Rubinstein, Meyer and Evans 2001). This process may result in several types of cognitive interference – “thoughts that intrude on task-related activity and serve to reduce the quality and level of performance (Sarason and Pierce 1996: 326).” First, lab studies find that when individuals switch between multiple tasks their working memory may become overloaded impairing performance (Rogers and Monsell 1995). This finding holds even when subjects did not expect to complete the prior task ever again (Allport and Wylie 2000). Second, lab studies also find that subjects have to apply cognitive resources to inhibit the stimuli from the prior tasks (Wylie and Allport 2000; Waszak, Hommel and Allport 2003). Finally, with greater cognitive interference individuals may be subject to higher levels of stress (Rogers and Monsell 1995; Monsell 2003) which may impair performance. These different factors suggest not only that specialization may dominate variety in the short-term, but also that the returns to specialization may be decreasing in variety since workers are not able to focus all of their cognitive resources on the present task. Therefore, we hypothesize:

HYPOTHESIS 2: *Same-day, different-task experience moderates the relationship between same-day, same-task experience and worker productivity, such that the positive effect of same-day, same-task experience is stronger under low same-day, different-task experience than under high same-day, different-task experience.*

This expectation runs counter to the findings of Schilling et al. (2003), that students in a lab for one day improved their performance most when playing two related games (i.e., not one game). The two studies differ in several key ways. First, the Schilling et al. study examined students who had not played a game before, while we studied workers repeating work day after day while they pursued their profession. Second, students were playing computer games, an activity considered fun by the participants. This task is very different from our procedural task context. Third, students were asked to complete their tasks within teams, while here we examined individual workers. Finally, the students were told to play at their own pace, and so had ample time for reflection, planning, and taking breaks, if desired. While we would expect a different outcome in a naturalistic setting with repetitive, procedural tasks combined with the theoretical arguments presented above, this question can and will be answered empirically in this paper.

2.2 Specialization and Variety in the Long Term

As detailed above, variety can affect performance both positively and negatively. While we hypothesize that specialization will dominate varied experience within a day, we posit that variety may prove more beneficial over many days. Before we specify the valuable effects of variety over time, we first discuss the expected relationship between experience and performance for workers completing repetitive, procedural tasks. With repetition, a worker gains knowledge and dexterity to help her complete the work more effectively. This performance improvement, resulting from the accumulation of experience, is traditionally captured in an experience curve (Newell and Rosenbloom 1981; Lapré 2011).

However, while skill may increase with experience, a worker's effort or motivation may decrease with experience. Over time, workers may grow bored with the job or suffer from burnout, both of which may lead to decreased effort (Pines and Maslach 1978; Staw 1980; Fisher 1993). Additionally, if boredom increases, workers may engage in distracting behaviors that raise their arousal levels but detract from performance (Vroom 1964; Scott 1966; Hackman 1969). Thus, given these joint effects, depicted in Figure 1, Staw (1980) posits that "most jobs have an inverted U performance curve simply because performance is generally a joint function of skills and effort" (p. 260).

***** FIGURE 1 ABOUT HERE*****

We test this relationship for a worker's experience curve. Prior research refers to the experience curve as both a progress function and a learning curve, among other terms (Yelle 1979; Dutton and Thomas 1984). We use the term *experience curve* in this paper to refer to changes in workers' productivity that are due to task experience, since the changes we observe in the data could result from learning or changing motivation. Therefore, our third hypothesis is:³

³ Prior research identifies U-shaped learning curves for relative performance measures such as organizational survival (Baum and Ingram 1998), profitability (Ingram and Simons 2002), and customer dissatisfaction (Lapré and Tsikriktsis 2006) due to competency traps (Levitt and March 1988). These authors note that they do not expect to

HYPOTHESIS 3: *Worker productivity follows an inverted U-shaped function for all prior days' same-task and different-task experiences.*

The next question we address is how variety affects productivity over time. First, variety may lead to learning benefits. Over time, a worker may identify learning opportunities across various task types. For example, a worker may recognize that a strategy used in one task can be used profitably in another area, or may realize that parts of strategies used in multiple tasks may be combined to yield a better performance outcome. Second, varied experience may provide motivational benefits. The negative effects of specialization likely become more salient as a worker completes more tasks and grows bored with the work. Therefore, with increasing variety, the worker may remain engaged with the work and thus exert more effort that may improve her performance over time or, at least, not degrade it. While additional setups from task change are still costly over time, these costs may be dominated by the benefits of variety.

How then will these effects manifest themselves in worker productivity? While specialization may provide greater benefits than will variety for lower levels of experience (Boh et al. 2007), the returns from specialization likely decrease at a faster rate than do the returns from varied experience, given the benefits outlined above. More interestingly, the learning and motivation benefits discussed earlier suggest that over time, specialization and variety may interact to improve performance. This idea is captured in the theoretical model of Lindbeck and Snower (2000), who argue for returns from task complementarities, suggesting that a worker's experience with "one task raises his productivity at another task" (p. 359). The complementarity may arise for learning reasons – e.g., varied experience may aid in learning how to learn (Ellis 1965); it may also help trigger a different learning process, in which discrepancies cause a worker to change her underlying theories about the process (Piaget 1963), resulting in performance improvement. Alternatively, the benefit may be motivational, as varied experience keeps a worker engaged, committed, and interested in her job so that she is willing to continue to take part in performance improvement efforts. While Lindbeck and Snower's model captures only one time period, these researchers note that returns from complementarities should "manifest themselves only with the passage of time" (p. 360). Given these reasons, we hypothesize the following:

HYPOTHESIS 4: *All prior days' different-task experience moderates the relationship between all prior days' same-task experience and worker productivity, such that the positive effect of all prior days' same-task experience is stronger under high all prior days' different-task experience than under low all prior days' different-task experience.*

find an *organizational*-level U-shaped curve for absolute measures such as completion time as used in this study. The combination of increasing skill and decreasing effort may create such curves at the *individual* level, however.

2.3 Examining the Effect of Variety: Cumulative Setups

Finally, we examine one specific potential mechanism through which variety can produce performance benefits. As a worker completes more tasks across different stages, she may see benefit not only in learning and motivation, but also in her ability to change tasks more effectively. While analytical (Tzur 1996) and case-based (Shingo 1989) studies in operations management have examined learning in setups, we are aware of no empirical work that has examined learning in workers' cognitive setups.

Although changeovers are costly, given the cognitive readjustments outlined above, it is possible that, with increasing experience in setups, a worker may minimize such costs. For example, a worker may learn how best to cope with cognitive interference; alternatively, she may gain a routinized skill in changeovers such that the steps are ingrained in her long-term, not just in her working memory. Consistent with this argument, research in psychology has found that, while individual task changes may be distracting, repeated practice in changing tasks may reduce cognitive switching costs (Gopher, Armony and Greenshpan 2000; Monsell 2003). In addition, given the discussion in Hypothesis 3 about the joint effects of learning and effort, the same forces likely are in play with respect to setups, so we would expect to see a U-shaped relationship with performance. Therefore, we hypothesize that:

HYPOTHESIS 5: *Worker productivity follows an inverted U-shaped function for cumulative task changes.*

3. Data and Empirical Strategy

3.1 Setting

We test our hypotheses using data from the home loan mortgage processing line at Shinsei Bank. By mid 2007, with the launch of a new home loan mortgage line, the process was structured as outlined in Figure 2. The parts of the process depicted by black text graphics in Figure 2 were completed automatically by computers and those parts of the process depicted by white text graphics were instead completed by human operators; the latter parts of the process serve as the focus for our analyses. Below we explain the process sequentially, for ease of understanding; however, an application did not necessarily proceed strictly in the manner described. Multiple parts of the process between decision points could and did run in parallel.

*****Insert Figure 2 about here *****

As Figure 2 shows, the process begins when an application is received and scanned. While the scanning does involve some human input (e.g., operator places application into machine), we do not include scanning in our analyses, in part because scanning is the only stage whose completion time is not captured at a sufficiently minute level. One operator might open the envelopes while another operator might place a stack of applications in the scanner. Additionally, scanning takes place in another part of

the building, done by individuals different than those who figure in the remaining stages.

After scanning, forms go to the custodian stage, during which a worker compares the scanned image to the received document and either accepts or rejects the scan. Any image rejected is returned to scanning, where the process begins again. After the custodian stage is completed, documents are tagged: subsections of each scanned document are marked electronically (tagged) for future processing. Next, the application data is “captured”: workers input data from the application into the computer system. Specifically, each worker sits at a computer equipped with two monitors. On one screen, the worker is presented with an image of the application; on the other screen, she enters relevant data in the appropriate fields. Separate parts of the application are entered in the Application capture 1 and Application capture 2 stages. During the subsequent preliminary information part of the process, workers record several data fields from the remainder of the application using the approach just described. Preliminary information 1 enters one set of data and Preliminary information 2 enters data from different images of the application.

After this work is completed, the inputs are compared to underwriting standards (within the computer system). If the application fails to meet these standards, an automatic rejection letter is sent. If the application passes, the computer checks to see if the application is complete. If the application is missing data, a request for more information is generated automatically, and when that arrives, the entire application is processed again. If no data is missing, the application proceeds to credit check. In the first stage of the credit check, a worker enters the data needed to request an external credit report. In the second stage, a worker types in relevant fields from a scan of the faxed credit report. The computer then compares the application again to underwriting standards, and if it passes, generates a request for more materials from the prospective borrower. The company also has call center operations to handle customers’ inbound questions and to make outbound calls encouraging submission of paperwork, but these are outside the scope of the present study, and thus are not included in the data analyzed here.

Once additional data is received and scanned, it proceeds through new custodian, document tagging, and two additional application-capture stages (all defined as separate stages, given that the work differs from the earlier stages with the same names). After being checked by computer for completeness and compared against underwriting standards, the file proceeds to the income tax stages. In the first stage, a worker submits a request to the Japanese tax authority for verification of income tax forms; in the second stage, the authority’s response is entered. The computer again checks the file against underwriting standards before progressing to the real estate stages, the first of which requests a real estate appraisal from an outside party, followed by the next stage, data entry. The final stage we analyze is credit approval, which is completed not by a specially trained credit expert, but by a line worker. This worker examines the application against a number of prespecified standards. The comparisons show on the computer as green when acceptably above standard, red when unacceptably below standard, or yellow

when marginal. If the application meets or fails to meet the standard, the worker approves or rejects it, accordingly. If the application is marginal or the worker believes that special circumstances exist, she can send it to a manager for further examination. We have no data on this further examination, so it is excluded from our analysis. Table 1 briefly describes the stages in the overall process.

*****Insert Table 1 about here *****

To summarize the structure of the Shinsei line, each of the seventeen stages analyzed is considered to be distinct. When a worker is assigned a task within a stage (e.g., Application capture 1), she completes all work for that stage, with no physical handoffs. When a worker completes a task, the system assigns a new task – a worker does not have an individual queue. Lunch and break times are noted within the system, and no tasks are assigned to workers during these times. The system provides workers no information about the state of the task queue; rather, each worker learns of the next task to complete when it arrives on her desktop. Line workers are not specialized to complete any given task (receiving no specialized training), including credit approval.

3.2 Data

Our sample includes all loan applications processed at Shinsei Bank between June 1, 2007, and December 30, 2009: 56,227 loan applications, totaling 601,788 individual stages completed by 140 workers. Twenty-nine workers in the data set appeared for fewer than 200 transactions each. All but five of these stayed at the firm for ten days or less, thus were either short-term temporary workers or had joined and immediately left the firm during a two-week probationary period. The remaining five individuals were managers who occasionally completed transactions when workers were absent. We dropped all these workers and their transactions from our analysis, leaving 598,393 transactions and 111 workers. At each process stage, Shinsei's IT system tracks detailed information on each loan application, from which we constructed the study's variables. We noted that employees were paid an hourly rate with no incentive pay, had no daily quota of tasks to complete, and were given no work performance targets. Additionally, pay raises were based on firm tenure, not performance. Management also reported that no workers were involuntarily separated from the company during the period of our research, outside of the two-week probationary period.

3.2.1 Dependent Variable.

Completion Time. We measure our dependent variable by calculating the number of minutes a worker took to complete the present stage and taking the natural log of the value to give the *completion time*. Processing time is a common measure for evaluating operational performance (e.g., Reagans, Argote and Brooks 2005). Shinsei management reported that faster processing time for individual tasks helped, in part, the company to more quickly process loan applications, helping the company compete more effectively by increasing the likelihood of securing customers. The mean of the unlogged variable is 2.74

minutes, and its standard deviation 3.54 minutes. Similar to the approaches used by Boh et al. (2007) and Narayanan et al. (2009), we run our analyses on all transaction data while controlling for the characteristics of each stage, permitting us to examine any given worker's complete work history at Shinsei during the time our data was collected by the bank.

3.2.2 Independent Variables. Choosing a formalization to operationalize variety is an important design consideration. At least two basic approaches can be employed: The first is to use a volume measure for both stage-specific experience and experience acquired at all other stages. The second is to use a share-based measure (such as a percentage) to examine the effect of differing types of prior experience. Here we use the former approach, both to be consistent with prior literature at the individual level (Boh et al. 2007) and because we are interested in task allocation at a micro level. In other words, we are concerned with whether an additional task should be allocated to specialized or varied experience, based on the amount of a worker's prior experience.

Stage-specific volume. To measure stage-specific (task-specific) experience, we construct variables that count the number of times an individual has executed the focal stage previously. We calculate both *same-day, stage-specific volume* and *all prior days' stage-specific volume*. For the same-day, stage-specific measure, we zero it out at the start of each day, then count the number of times a worker executes that stage on a particular day, prior to execution of the current task. The all prior days' stage-specific measure counts the number of times an individual has executed the focal stage prior to the start of the current day. Thus, while same-day, stage-specific volume changes throughout a day, all prior days' stage-specific volume does not.

Other-stage volume. We also calculate similar measures for each worker's other experience. First, we calculate *same-day, other-stages' volume*. This measure, zeroed out at the start of each day, captures the number of times a worker executes *all other* stages on a given day. Next, we calculate *all prior days' other-stages' volume*. This variable counts the number of times an individual has executed *any other* stages prior to the start of the current day.

Given the theory motivating Hypothesis 3, our models include the quadratic variable for both all-prior-day-volume measures, but not quadratic terms for the same-day volume measures. We note that if we include the quadratic terms for same-day volume, each term becomes significant, but negative returns do not occur until the 99th percentile of the distribution. Therefore, we exclude quadratic variables from the analyses, but note that all hypotheses hold with them included.

Cumulative stage changes. We calculate the number of times a worker had changed stages prior to the present task. We include both a linear and a quadratic measure for the volume of stage changes. Since we are interested in whether cumulative stage changes reduce completion times, when a stage change occurs, we interact these two variables with the indicator for a stage change.

We do not log our experience measures, as we use the exponential form for our experience-curve analyses. We use the exponential form for two reasons. First, while the exponential form is derived from theory and supported empirically, the power form (log-log) comes simply from empirical observation (Levy 1965; Lapré et al. 2000). Second, as Lapré and Tsikriktsis (2006) note, if any experience has been gained prior to the start of data collection, then the power form will be biased. While our data captures the start of the entire IT-enabled work process, some individual stations came online before June 2007; thus, some workers had acquired experience prior to the initial data-collection point of our study.

3.2.3 Control Variables. Table 2 details the control variables included in the analyses.

*****Insert Table 2 about here *****

One final point: we have no data regarding characteristics of individual loans (e.g., loan amount), and so include no controls for these factors. Two reasons lead us to believe our results are robust without these controls. First, according to Shinsei personnel, differences in borrowers or loan sizes do not affect loan processing, just the credit decision. Second, and more importantly, loans within a stage are assigned randomly. The IT system presents a task to a worker when she finishes her prior task, without regard to loan characteristics. We provide more details on this process below (see section 3.4).

Table 3 provides summary statistics for the variables used in our models.

*****Insert Table 3 about here *****

3.3 Empirical Approach

We wish to estimate models that capture the effects of specialized and varied experiences on task-level performance. Since our data is a complete history of each individual's work volume over three years, we need to select a model that accounts for autocorrelation, contemporaneous correlation, and heteroskedasticity (Beck 2001; Lapré and Tsikriktsis 2006). We thus chose to use Prais-Winsten regression, as detailed by Lapré and Tsikriktsis (2006), i.e., with panel-corrected standard errors adjusted for heteroskedasticity and panel-wide, first-order autocorrelation (Stata command: `xtpcse`).

In our primary analyses, we first estimate a model with just control variables. Next we add linear and quadratic variables to capture the total *cumulative volume* for each worker, then separate volume into *same-day cumulative volume* and *all prior days' cumulative volume*. To test Hypotheses 1 and 3, we further divide experience into stage-specific and other-stage volumes, creating four variables: *same-day, stage-specific volume*; *same-day, other-stage volume*; *all prior days' stage-specific volume*; and *all prior days' other-stage volume* (with linear and quadratic terms for the last two experience types). We then test Hypotheses 2 and 4 by adding the interaction terms for *same-day, stage-specific volume* \times *same-day, other-stage volume* and *all prior days' stage-specific volume* \times *all prior days' other-stage volume*. Finally, to test Hypothesis 5, we add the interaction of the *stage change* indicator with the linear and quadratic terms for *cumulative stage changes*.

Our dependent variable is the log of the completion time for task k in stage i completed by individual j . Our complete model (Column 7 in Table 4a) is the following:

$\ln(\text{Completion Time})_{ijk} =$

$$\begin{aligned} & \beta_0 + \beta_1 \text{Same-Day, Stage-Specific Volume}_{ijk} + \beta_2 \text{Same-Day, Other Stages' Volume}_{ijk} + \beta_3 \text{Same-Day,} \\ & \text{Stage-Specific Volume}_{ijk} \times \text{Same-Day, Other Stages' Volume}_{ijk} + \\ & \beta_4 \text{All Prior Days' Stage-Specific Volume}_{ijk} + \beta_5 \text{All Prior Days' Stage-Specific Volume}_{ijk}^2 + \\ & \beta_6 \text{All Prior Days' Other Stages' Volume}_{ijk} + \beta_7 \text{All Prior Days' Other Stages' Volume}_{ijk}^2 + \\ & \beta_8 \text{All Prior Days' Stage-Specific Volume}_{ijk} \times \text{All Prior Days' Other Stages' Volume}_{ijk} + \\ & \beta_9 \text{Stage Change}_{ijk} + \beta_{10} \text{Stage Change}_{ijk} \times \text{Cumulative Stage Changes}_{ijk} + \\ & \beta_{11} \text{Stage Change}_{ijk} \times \text{Cumulative Stage Changes}_{ijk}^2 + \beta_{12} X_{ijk} + \lambda_t + u_{ijk} \end{aligned}$$

where X_{ijk} is a vector of the individual-task control variables and λ_t is a year indicator to control for unobserved factors that could affect the average trend in completion time.

Considering our hypotheses, Hypothesis 1 predicts that over the course of a day, variety will lead to an increase in completion time compared to specialized experience: $\beta_1 < \beta_2$. Note that, since lower completion time means improved performance, a negative coefficient here corresponds to a “better” outcome. Similarly, when we predict an inverse U-shaped relationship with productivity, since our dependent variable is completion time this means that we will test for a U-shaped relationship with completion time. Hypothesis 2 predicts that same-day, stage-specific and same-day, other-stage volumes will interact to worsen performance: $\beta_3 > 0$. Hypothesis 3 predicts that experience over time will exhibit a U-shaped relationship with completion time: $\beta_4 < 0$ and $\beta_5 > 0$, and $\beta_6 < 0$ and $\beta_7 > 0$. Hypothesis 4 predicts that the interaction of all prior days’ stage-specific volume and all prior days’ other-stage volume will be related to improved productivity: $\beta_8 < 0$. Finally, Hypothesis 5 predicts that setup times during changeovers will exhibit a U-shaped relationship with completion time: $\beta_{10} < 0$ and $\beta_{11} > 0$.

3.4 Data-Generation Process

An important question arises about our study’s underlying data-generation process. The concern is that variety might be assigned to the best (or worst) workers and therefore any productivity effects from variety may be due to innate characteristics of the workers, not the variety itself. As described above, however, when Shinsei management redesigned the home loan mortgage processing line, their goal was to remove the human element as much as possible. In describing the redesign, one Shinsei senior manager noted, “When the machines orchestrate the work, the people can just be plugged in.”⁴

⁴ Interview conducted with Shinsei senior manager on February 15, 2011.

Therefore, at Shinsei Bank, management reported that task assignment was structured through the system and variety was not used to motivate workers. Rather, workers were assigned to a given stage at the start of each day. The system did this automatically, although it was likely to keep workers at their last station, if possible (the departure or absence of a worker at another station could lead to a change, however). Then during the day, if the system identified a backup at a given station, it would assign the next available worker to that station. Depending on demand dynamics, the system could reallocate workers repeatedly over the course of a day. Management reported that workers did not request, and were not given, additional variety in task assignment. Thus, Shinsei managers noted that line managers did not reallocate work or prioritize more talented individuals to receive variety, but rather the system simply used the aforementioned algorithms to assign work.

We note that this process will result in a positive correlation between variety and cumulative volume. Individuals working longer are at increased risk of receiving tasks from a different stage due to a backup in any area. In fact, all prior days' stage-specific volume and all prior days' other-stages' volume are positively correlated (correlation coefficient = 0.22, $p < 0.01$). As long as this assignment occurs as reported, the correlation between measures does not bias our results.

As an additional check on variety assignment, we examined whether, on average, individuals experienced differential task variety within a day. If management were allocating variety in tasks based on worker characteristics, we would expect to see persistent variation across our sample. Therefore, we regressed the worker indicator variables and an indicator for each day on workers' overall daily variety (i.e., $\text{Daily Task Variety}_{it} = \text{Worker Indicator}_i + \text{Day Indicator}_d$). Since this model takes place at the daily level (not at the task level), we calculated daily variety by constructing a Blau measure using each day's realized volume.⁵ After running this model using Prais-Winsten regression, we conducted a χ^2 test on the individual indicators to find the p-value is not significant ($p = 0.87$, $\chi^2(110) = 93.52$). The fact that the indicators as a group are not significant and only seven of the individual indicators are significant (out of 110, one would expect approximately six false positives for $p < 0.05$) increases our confidence that task variety is assigned to workers along the lines described by management.

4. Results

Table 4a summarizes results from the regression of completion time on experience. In Column 1, we include all control variables, while in Column 2 we see support for an overall U-shaped experience curve. In Column 3, we find evidence for within-day specificity of learning. Both linear experience

⁵ A Herfindahl index is calculated by identifying the percentage of an individual's total daily experience that is represented by each stage, then squaring that value and summing the components. However, since a larger value for the Herfindahl index is related to increased specialization, we subtract the index value from one. The result, daily task variety, is also known as the Blau index (Harrison and Klein 2007).

variables are negative and significant, while the quadratic term for all prior days' cumulative volume is positive and significant. The coefficient for the linear term for the same-day cumulative volume variable is more negative than the linear coefficient for all prior days' cumulative volume (χ^2 test: 89.58, $p < 0.001$).

Turning to Column 4 and our tests of Hypotheses 1 and 3, we see that both within-day experience variables are negative and significant. Each additional unit of same-day, stage-specific volume decreases completion time by approximately 0.034%, or 5.6 seconds for a transaction of average length, while each additional unit of same-day, other-stages' volume improves performance by approximately 0.012%, or 2.0 seconds for a transaction of average length. As Hypothesis 1 predicts, the coefficient on same-day, stage-specific volume is more negative than the coefficient on same-day, other-stages' volume, and the difference is significant (χ^2 test: 174.00, $p < 0.001$). We also find that both types of all prior days' volume have a U-shaped relationship with completion time, supporting Hypothesis 3. Both curves turn negative within the support of the distribution – the inflection points are at 1.7 and 1.8 standard deviations above the mean for all prior days' stage-specific volume and all prior days' other stages' volume, respectively. Examining the coefficients, we find that for all prior days' stage-specific volume, the linear term is more negative and the quadratic term more positive than are the comparable terms for all prior days' other stages. In other words, specialized experience is related to better performance initially than that achieved by varied experience, but the benefit of specialized experience degrades more rapidly.

*****INSERT TABLES 4A & 4B ABOUT HERE*****

Moving to Column 5, we examine the coefficient for the interaction of same-day, stage-specific volume and same-day, other-stages' volume and find that it to be positive, providing support for Hypothesis 2. In other words, increasing an individual's variety of experience over the course of a day decreases the marginal benefit of each subsequent task's being executed on the individual's task productivity. Next, we see that the coefficient on the interaction of all prior days' stage-specific volume and all prior days' other-stages' volume is negative and significant, providing support for Hypothesis 4. To examine the interaction further, we plot the net effect of the interaction (main effects added to the interaction terms for multiple values of experience, see Figure 3a) and find that the plot supports the view that varied experience is related to ongoing performance improvement. In Column 6, we include all four possible interaction terms for the all prior days' volume measures (i.e., interacting all linear and quadratic terms for the all prior days' volume measures). Figure 3b plots these values. As the figure shows, varied experience eventually is related to performance superior to that achieved with specialization.

***** FIGURES 3a & 3b ABOUT HERE*****

Finally, in Column 7 we examine whether learning occurs during changeovers, as predicted by Hypothesis 5. We note that the main effect of stage change is higher average completion time, providing support for the laboratory findings of Schultz et al. (2003). Information workers switching from one task

to the next must engage in a “cognitive setup,” slowing productivity. Adding an indicator for the second task, post-stage change shows that performance is typically worse than average, although there is a statistically significant improvement as compared to the first task. When we include the interactions of cumulative stage changes with the indicator stage change, we find support for Hypothesis 5. Thus, controlling for a worker’s specialized and varied experience, we find evidence for a U-shaped relationship with performance.

We note that, in addition to our hypotheses, several other coefficients in the model are of interest, as shown in Table 4b. First, consistent with KC and Terwiesch (2009), we find that increasing the *load* on workers during a shift is related to decreased processing times. However, these gains do not appear to be sustainable, as worker *overwork* is related to increased processing time. Second, we see that higher levels of monthly *utilization* are related to decreased processing time. Thus, by including these variables in our models, not only do we control for factors that may influence our results, but also we are able to replicate findings of KC and Terwiesch (2009) in a non-healthcare setting, showing that service rates are endogenous to the load on a system (Schultz et al. 1998).

Next, by examining the coefficient values for the day indicators, we find that, on average, completion times are slower on Monday and faster on Saturday than on any other day of the week. On Saturday, work volumes are lower and workers are sent home when the day’s work is completed, so workers might be especially eager to finish their work quickly, to leave as early as possible. This suggests that not only do incentives work in this context, although the company does not use monetary incentives to encourage faster completion time, but also slack in the system exists, since workers can complete their tasks more quickly than on weekdays without negatively impacting quality. Finally, we note that the variable for the year 2009 is positive and significant, indicating that, holding all other variables constant, completion times in 2009 were slower than in 2007 or 2008. 2009 was an exceptional year in the global financial markets due to the liquidity crisis around the world that restricted lending. When asked, a Shinsei manager speculated that this decrease in productivity in 2009 was due to the lower volume of applications as well as the general distractions and uncertainty felt by staff due to the crisis. While the manager’s first point should be largely covered in our models by their inclusion of load and utilization variables, the latter point could lead to the decreased productivity. For example, prior work suggests that productivity may suffer from external uncertainties and any threat of layoffs (Greenhalgh and Rosenblatt 1984), even though we note that Shinsei did not lay off any workers.

4.1 Robustness Checks

To further examine our results’ robustness, we explore several additional factors (results not shown). First, we repeat our analyses, excluding those for 2009, finding the results continue to support our hypotheses. Second, one can consider additional controls for variety. Narayanan et al. (2009) include

a Herfindahl-based measure for variety. When included with variables for both specialized and other experience, this Herfindahl variable effectively captures how the *other* experience is distributed. Therefore, we add to our model a Blau measure for all prior days' variety that is calculated the same way as the total variety measure discussed in Footnote 5, except it is at the task level, and thus is updated after each executed task. We also include the interaction of this variable with all prior days' stage-specific volume to capture any additional complementarities between variety and specialized experience. Including these variables does not change the support for our hypotheses. Additionally, while the coefficient on all prior days' variety is positive ($\beta=0.2571$, $p<0.001$), suggesting that experience distributed across more categories hampers productivity on an absolute basis, the *interaction* term is negative ($\beta=-3.921e-05$, $p<0.001$), suggesting that such variety provides additional marginal value for each unit of specialized volume.

Next, given that the average worker has thousands of observations, there is a concern that the Prais-Winsten panel-corrected standard errors may not adequately account for longer-term autocorrelation. Therefore, we repeat our analyses using fixed-effects regression models with block-bootstrapped standard errors (Stata command `xtreg, vce(bootstrap)`)—and continue to find support for our hypotheses. Also, since individuals learn at different rates, we repeat the analyses using a mixed-effects model that permits the experience variables and their interaction to vary for each individual (Stata command `xtmixed`). This approach again shows support for the study's hypotheses. Finally, given that many workers executed tasks across many stages in our data, there is concern that standard errors might differ across both workers and stages. Therefore, we repeat our analyses using ordinary least squares regression to cluster standard errors by worker and stage (Stata command `cluster2`, Cameron, Gelbach and Miller 2010), and again find our hypotheses supported. Additionally, it is possible that our standard errors could differ across workers and applications; we therefore cluster the standard errors by worker and application—yet again finding our hypotheses supported.

4.2 Limitations and Venues for Future Research

Although we explored several explanations of our findings' support for our hypotheses, our investigation is subject to limitations. First, any non-random assignment of variety to individuals could bias our results. While discussions with Shinsei management give us confidence that our results are properly identified, future work could implement a field experiment to further examine our findings. Second, due to factors such as the company's IT system and the nature of the focal task, quality is high in this context and shows little variation. While our results are significant both statistically and organizationally, future work could examine the effect of these variables on quality performance and other factors such as total factor productivity and workers' creativity and innovation. Third, our analysis' examination of work *variety* raises the question of how related the tasks are that we study. Boh et al.

(2007) and Narayanan et al. (2009) use a software module, while Schilling et al. (2003) use the concept of a game to define “relatedness.” In this study, we treat all the work as related, since it all relates to a single product: a home loan mortgage. Tasks in the present context also are all related, as they involve analyzing and inputting data into a computer. Future work could seek to identify the dimensions of relatedness and examine the effects on work when tasks differ increasingly one from another.

Fourth, the present study examines our hypotheses in a procedural task setting similar to many operational contexts in which workers exert physical and mental effort. Future work could explore how findings differ in a purer knowledge-based work setting. Fifth, the present study uses a single day and multiple days to represent short and long time periods, respectively. While these are appropriate for the present context, future work could examine different contexts with varying task lengths. Finally, the present study examines only one organization, an undesirable but necessary consequence of both gaining access to such detailed data and learning the intricacies of their context. While we believe the theory in our work holds true in other contexts, future work could rigorously examine the validity of our hypotheses elsewhere.

5. Discussion and Conclusion

In most contexts that involve repetitive work, managers have an important decision to make: how to assign tasks to workers? While some scholars argue for specialization (Boh et al. 2007), others recommend varying the assignment (Hackman and Oldham 1976; Narayanan et al. 2009). Our findings suggest the answer to this question is contingent. While a specialized assignment strategy is related to improved productivity during the day (i.e., in the short term), variety is related to improved productivity over time (i.e., in the longer term). This study’s main finding suggests that, in contexts characterized by repetitive work, managers should consider keeping workers specialized on a task over the course of a day, while varying their task assignments over time.

From a strategic perspective, the following question can be posed: what size gain might a manager achieve if she were to play the specialization–variety game strategically? For simplicity’s sake, let’s assume a worker typically completes one hundred tasks in a day. Using the coefficients from Table 4a, Column 6, we compare the productivity difference for a worker under a specialized (all one task) strategy and a varied strategy (four stages completed: 25, 25, 25, 25). Focusing on just contributions from the same-day experience variables (holding all other variables constant), a worker completing just one task would complete the 100th task approximately 3.7% faster than average, while a worker completing four different stages would complete the 100th task approximately 2.3% faster than average – an absolute advantage of 1.4% for the specialized strategy, compared to the varied strategy.

Looking at results over time tells a different story. Assuming a worker completes 100 transactions

per day for a total of 10,000 transactions (approximately one standard deviation above the mean), one can consider the overall, experience-based differences across a specialized or a varied task strategy (again, for the latter, each of four stages completed 25, 25, 25, 25 per day). Holding all variables constant except those concerning experience and task-change, we find that a worker following a varied strategy could complete her work 23% faster than a worker following a specialized strategy. Thus, our results suggest that careful task assignment—keeping workers focused on a specialized task during the day each day, but varying task assignment across days—might improve operational performance.

We note that while our findings are both statistically and organizationally significant, our models show that a meaningful portion of the variance in our setting remains unexplained (the R^2 in our final model is 34%). Future work could examine the role other factors play in explaining additional variance. First, while we investigate completion time at the task-level, analyzing the underlying parts of an overall task (i.e., sub-tasks) may offer further insight on how to structure tasks for learning (Kantor and Zangwill 1991; KC and Staats 2011). Second, prior work finds that repeated interactions between individuals can affect performance (Huckman, Staats and Upton 2009; Schultz, Schoenherr and Nembhard 2010; Huckman and Staats 2011). Future research could examine how peer effects interact with variety, if at all. By examining whether co-workers are located next to one another or whether they share breaks may reveal interaction effects with variety and a direct effect on performance. Finally, research finds that external conditions (such as the weather) affect decision making (Simonsohn 2011). Extending this work to operational settings may provide further insight on worker motivation and performance.

The current investigation provides insight into how work can be structured effectively across tasks and over time. Thus, it responds to calls from the scholarly literature for more nuanced theorizing on experience and productivity (Gino et al. 2010; Lapré and Nembhard 2010; Argote and Miron-Spektor 2011). First, with our detailed data on repetitive, procedural tasks, we provide evidence for inverse U-shaped experience curves for workers, likely due to joint effects of skill and motivation (Staw 1980).

Second, our study builds on recent research on specialization and variety at the individual level (Boh et al. 2007; Narayanan et al. 2009) by examining the topic outside the software maintenance environment. Resolving a software bug took, on average, two and a half *days* in the prior studies; here we examine a repetitive task that took a given worker, on average, two and a half *minutes* to perform. Workers therefore can execute more tasks over time, and their risk of boredom is likely higher in our study. Also, the data entry captured in our study requires less specialized skill than does debugging software code. Thus, our investigation examines highly repetitive, procedural tasks that both differ from software development and represent many different operational contexts.

Third, we inject the temporal dimension into the debate of specialization and variety for individual workers, looking both within a day and across days. By separately theorizing about and then

evaluating the effect of each strategy over the course of a day and over many days, we are able to separate when and how each strategy relates to improved performance. Introducing a temporal dimension to the specialization-versus-variety debate also allows us to examine how the two strategies should be balanced, a need previously identified by Narayanan et al. (2009). Our finding of these strategies' complementarities highlights the fact that balancing the two strategies is not the only matter; since the two strategies are related, ways must be found to turn them into mutually reinforcing strategies.

Fourth, we consider whether the effect of varied experience on productivity is due to the direct effect of varied experience or from the interaction of varied experience and specialized experience. In so doing, we are the first to unpack these benefits at the level of the individual (see, Schilling et al. 2003; Clark and Huckman 2011, for analyses at the team and organizational level, respectively). Within a day, we find a residual cost to variety as the returns to specialization decrease in the amount of variety. This finding brings work on task-change paradigms out of the lab and into the field. While variety is distracting during a day, we find support for complementarity over time, as variety can aid both learning and motivation. Altogether, understanding the mechanisms by which variety helps (or hurts) performance creates the ability to theorize more effectively and provide more useful managerial advice.

Fifth, we examine one mechanism through which variety can improve performance: cumulative setups. While prior work finds that setups due to changing tasks are costly (Schultz et al. 2003), here we find that increasing cumulative experience for workers helps to mitigate these costs. Finally, our results contribute also to the development of behavioral theory in operations (Boudreau et al. 2003; Bendoly, Donohue and Schultz 2006; Gino and Pisano 2008). Our model integrates the operations management and organizational behavior perspectives; thus, with a finer-grained understanding of the relationship between experience and performance, better operating systems can be designed.

Our study's implications for operational performance offer opportunities for managerial action, given the increasing fragmentation of work. Advances in information and communication technologies permit organizations to divide work into very small tasks for distribution to workers who may or may not be collocated (c.f., Levy and Murnane 2004). At Shinsei, the company redesigned their process with the objective of removing human variability, instead having the IT system control the process. While the gains from IT may be valuable, this study highlights that task assignment still plays an important role in determining worker productivity; we find that behavioral elements still impact completion time. Although ongoing advances in technology may create opportunities to establish virtual factories (Stross 2010), our results highlight the need to identify and then implement algorithms for task assignment in such contexts, and in operations more generally that consider the gains and costs from both specialization and variety.

Altogether, our results highlight that in task assignment, the important relationship to examine is not specialization *versus* variety, but specialization *and* variety.

Table 1. Description of stages analyzed.

Name	Separate Stages	Description
Custodian	2	Scans and actual documents are compared to confirm quality (initial and additional data).
Doc tagging	2	Images on scanned documents are tagged for data entry (initial and additional data).
Application capture	4	Data from application forms are entered into computer (divided into two stages for different forms, for initial and additional application capture).
Preliminary information	2	Specific fields of data from additional forms are entered into computer (divided into two stages, corresponding to different forms).
Credit check	2	Stage 1 requests a credit report; stage 2 enters data from the report.
Income tax	2	Stage 1 requests tax verification data; stage 2 enters data from report.
Real estate	2	Stage 1 requests a real estate appraisal; stage 2 enters data from appraisal.
Credit approval	1	Application is accepted, rejected, or routed to an expert, based on underwriting criteria.

Table 2. Control variables.

Name	Description
Load	Prior studies find that workers increase their processing speed as the load on a system increases (Schultz et al. 1998; KC and Terwiesch 2009). Following KC and Terwiesch (2009), we construct a variable, <i>load</i> , that measures the percentage of workers who completed transactions during the hour in which the focal task started.
Overwork	While increasing system load may be related to decreased processing time, if overload continues too long, worker performance may be negatively impacted (KC and Terwiesch 2009). Thus, we construct a variable, <i>overwork</i> , to control for this effect. Overwork is calculated for each worker and each transaction as follows: $overwork_{i,K} = \frac{1}{N(K,i)} \sum_{j=i-N(K,i)}^{i-1} (Load_j - \overline{Load}_{s(j)})$. $N(K,i)$ is a count of the transaction requests throughout the prior K periods up to $t(i)$, the time when task i arrives, while $\overline{Load}_{s(j)}$ captures the average load for shift s . The K periods are measured in hours, and as in KC and Terwiesch (2009), $K = 4$.
Utilization	Workers lack queue awareness, but managers can view system backlog. While managers do not reallocate volume based on backlog, managers could possibly encourage workers to work faster. Also, a higher backlog decreases the likelihood that the system will allocate a different stage to a worker. Thus, we control for system utilization on a monthly basis, dividing total minutes that workers were working by total minutes available to work (shift length minus lunch and breaks), for the <i>prior</i> thirty days. For the first month, we calculate utilization for all prior days, setting the value to zero for the first day.
Defect	At Shinsei, two workers complete data-entry tasks, and their outputs are compared. If a discrepancy appears, the work is given to two other workers. This process repeats until two workers' output agrees. Therefore, we construct an indicator variable, <i>defect</i> , that equals one if an output was rejected or zero otherwise.
Stage change	We construct an indicator variable set to one when a stage change occurs (when a worker switches from completing work in one stage to doing so in another stage during the same workday), otherwise this variable is set to zero. Workers who change stages do not change physical stations.
2 nd Task, post-stage change	We construct an indicator variable set to one for the second task after a stage change (at the same stage in the same workday), otherwise this variable is set to zero.
Day-of-week	To control for day-of-week effects (Bryson and Forth 2007; Anbalagan and Vouk 2009; Schultz et al. 2010), we construct indicators for Tuesday through Saturday (Monday is the missing category). Work during the week is from 9:00am to 6:00pm. On Saturday, work begins at 9:00am, and ends when the work is finished. Realized volume for Monday through Saturday is 21%, 18%, 20%, 19%, 18%, and 4%, respectively.
Year indicators	We add indicators for the year each task was completed (with 2007 as the excluded category). This variable controls for any environmental differences across time.
Stage indicators	In order to compare performance across stages, we control for stage differences by including indicators for all but one of the 17 stages that appear in the data.
Individual indicators	To control for time-invariant aspects of workers, such as innate skill, we include indicators set to one when a worker completes a task and zero otherwise. All productivity hypotheses are tested "within-worker."

Table 3. Summary Statistics for Productivity Analysis ($n = 598,393$).

Variable	Mean	σ	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Log completion time	0.39	1.15														
2. Cumulative volume	5,470	4,620	-0.07													
3. Same day cumulative volume	110.8	119.7	-0.19	0.21												
4. All prior days' cumulative volume	5,295	4,527	-0.04	0.96	0.16											
5. Same day stage-specific volume	62.0	73.7	-0.20	0.12	0.46	0.08										
6. Same day other stage volume	61.7	92.5	-0.12	0.19	0.87	0.15	0.02									
7. All prior days' stage-specific volume	2,326	2,362	-0.15	0.70	0.18	0.62	0.24	0.09								
8. All prior days' other stage volume	2,977	3,302	0.01	0.87	0.16	0.87	0.00	0.20	0.25							
9. Load	0.64	0.23	-0.10	-0.05	0.19	-0.07	0.19	0.12	-0.02	-0.05						
10. Overwork	0.01	0.21	-0.04	0.00	0.02	0.00	0.02	0.02	-0.01	0.01	0.70					
11. Utilization	0.52	0.06	0.03	0.02	-0.13	0.03	-0.11	-0.10	0.02	0.01	-0.07	0.01				
12. Stage change	0.07	0.26	0.04	0.07	0.01	0.08	-0.06	0.05	-0.03	0.12	-0.05	0.00	0.01			
13. 2nd Task, Post-Stage Change	0.05	0.21	0.02	0.05	0.02	0.05	-0.04	0.04	-0.01	0.07	-0.02	0.01	0.00	-0.06		
14. Cumulative stage changes	404.3	486.1	-0.05	0.86	0.26	0.87	0.12	0.25	0.48	0.83	-0.03	0.01	0.00	0.11	0.07	
15. Defect	0.03	0.16	0.06	-0.06	0.02	-0.06	0.06	-0.01	0.00	-0.08	0.02	-0.01	0.00	-0.02	-0.02	-0.05

Note. Bold denotes significance of less than 5%.

Table 4a. Summary regression results on completion time of experience ($n = 598,393$).

	Dependent Variable: Log Completion Time							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Cumulative volume		-2.464e-05*** (1.203e-06)						
Cumulative volume ²		9.241e-10*** (5.342e-11)						
Same-day, cumulative volume			-1.143e-04*** (9.361e-06)					
All prior days' cumulative volume			-2.519e-05*** (1.211e-06)					
All prior days' cumulative volume ²			9.483e-10*** (5.457e-11)					
Same-day, stage-specific volume				-3.426e-04*** (1.616e-05)	-3.980e-04*** (1.829e-05)	-3.725e-04*** (1.829e-05)	-3.976e-04*** (1.829e-05)	Hypothesis 1
Same-day, other stages' volume				-1.243e-04*** (1.120e-05)	-2.100e-04*** (1.507e-05)	-2.062e-04*** (1.504e-05)	-2.097e-04*** (1.507e-05)	
Same-day, stage-specific volume × Same-day, other stages' volume					9.790e-07*** (1.005e-07)	9.738e-07*** (1.002e-07)	9.770e-07*** (1.006e-07)	Hypothesis 2
All prior days' stage-specific volume				-5.141e-05*** (1.746e-06)	-4.082e-05*** (1.944e-06)	-1.446e-04*** (4.741e-06)	-4.019e-05*** (1.946e-06)	Hypothesis 3
All prior days' stage-specific volume ²				4.057e-09*** (1.266e-10)	4.727e-09*** (1.357e-10)	2.287e-08*** (7.230e-10)	4.696e-09*** (1.357e-10)	
All prior days' other stages' volume				-1.781e-05*** (1.754e-06)	-9.512e-06*** (1.772e-06)	-1.327e-05*** (2.625e-06)	-8.671e-06*** (1.782e-06)	Hypothesis 4
All prior days' other stages' volume ²				1.005e-09*** (1.048e-10)	9.782e-10*** (1.047e-10)	6.454e-10*** (1.805e-10)	9.629e-10*** (1.054e-10)	
All prior days' stage-specific volume × All prior days' other stages' volume					-3.291e-09*** (2.355e-10)	1.401e-08*** (1.715e-09)	-3.334e-09*** (2.361e-10)	
All prior days' stage-specific volume × All prior days' other stages' volume ²						-4.263e-13** (1.422e-13)		
All prior days' stage-specific volume ² × All prior days' other stages' volume						-3.620e-12*** (2.351e-13)		
All prior days' stage-specific volume ² × All prior days' other stages' volume ²						1.364e-16*** (1.991e-17)		
Stage change × Cumulative stage changes							-1.108e-04*** (1.917e-05)	Hypothesis 5
Stage change × Cumulative stage changes ²							3.778e-08*** (9.363e-09)	
Stage change	0.0943*** (0.0037)	0.0951*** (0.0037)	0.0951*** (0.0037)	0.0899*** (0.0037)	0.0895*** (0.0037)	0.0890*** (0.0037)	0.1293*** (0.0070)	
2nd task, post-stage Change	0.0611*** (0.0045)	0.0618*** (0.0045)	0.0618*** (0.0045)	0.0575*** (0.0045)	0.0573*** (0.0045)	0.0567*** (0.0045)	0.0572*** (0.0045)	
Stage indicators	Significant	Significant	Significant	Significant	Significant	Significant	Significant	
Individual indicators	Significant	Significant	Significant	Significant	Significant	Significant	Significant	
Observations	598,393	598,393	598,393	598,393	598,393	598,393	598,393	
Number of Individuals	111	111	111	111	111	111	111	
R-Squared	0.3348	0.3360	0.3365	0.3388	0.3391	0.3428	0.3394	
ρ	0.2794	0.2779	0.2772	0.2747	0.2745	0.2696	0.2740	
Wald Chi-Squared	361150***	356704***	357022***	358494***	353765***	355890***	354287***	

Notes. *, ** and *** denote significance at the 5%, 1% and 0.1% levels, respectively. Prais-Winsten regression models with panel-corrected standard errors adjusted for heteroskedasticity and panel-wide first-order autocorrelation. Results are shown in Table 4b for the indicators for load, overwork, utilization, defect, day of week, and year.

Table 4b. Additional coefficients from models in Table 4a ($n=598,393$).

	Dependent Variable: Log Completion Time						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Load	-0.3786*** (0.0097)	-0.4159*** (0.0099)	-0.4047*** (0.0100)	-0.3896*** (0.0100)	-0.3727*** (0.0100)	-0.3607*** (0.0100)	-0.3728*** (0.0100)
Overwork	0.2314*** (0.0095)	0.2591*** (0.0096)	0.2509*** (0.0097)	0.2396*** (0.0097)	0.2262*** (0.0097)	0.2157*** (0.0097)	0.2263*** (0.0097)
Utilization	-0.0832** (0.0262)	-0.1041*** (0.0262)	-0.1254*** (0.0263)	-0.1399*** (0.0263)	-0.1385*** (0.0263)	-0.1492*** (0.0261)	-0.1379*** (0.0262)
Defect	0.3072*** (0.0067)	0.3038*** (0.0067)	0.3038*** (0.0067)	0.3037*** (0.0067)	0.3040*** (0.0067)	0.3033*** (0.0067)	0.3041*** (0.0067)
Tuesday	-0.0075 (0.0047)	-0.0112* (0.0047)	-0.0118* (0.0046)	-0.0130** (0.0046)	-0.0122** (0.0046)	-0.0102* (0.0046)	-0.0123** (0.0046)
Wednesday	-0.0117** (0.0045)	-0.0133** (0.0045)	-0.0137** (0.0045)	-0.0148*** (0.0045)	-0.0147** (0.0045)	-0.0127** (0.0045)	-0.0147** (0.0045)
Thursday	-0.0139** (0.0047)	-0.0172*** (0.0047)	-0.0180*** (0.0047)	-0.0197*** (0.0047)	-0.0188*** (0.0047)	-0.0166*** (0.0047)	-0.0188*** (0.0047)
Friday	-0.0135** (0.0047)	-0.0189*** (0.0047)	-0.0197*** (0.0047)	-0.0222*** (0.0047)	-0.0206*** (0.0047)	-0.0182*** (0.0047)	-0.0207*** (0.0047)
Saturday	-0.1500*** (0.0077)	-0.1546*** (0.0077)	-0.1601*** (0.0077)	-0.1677*** (0.0077)	-0.1680*** (0.0077)	-0.1681*** (0.0077)	-0.1675*** (0.0077)
Year 2008	-0.0348*** (0.0042)	0.0353*** (0.0060)	0.0399*** (0.0060)	0.0617*** (0.0059)	0.0441*** (0.0060)	0.0630*** (0.0060)	0.0430*** (0.0060)
Year 2009	0.1319*** (0.0072)	0.2448*** (0.0103)	0.2500*** (0.0102)	0.2553*** (0.0102)	0.2280*** (0.0102)	0.2170*** (0.0102)	0.2279*** (0.0102)
Constant	-2.1961*** (0.0694)	-2.1663*** (0.0692)	-2.1622*** (0.0691)	-2.1702*** (0.0688)	-2.1686*** (0.0688)	-2.1527*** (0.0685)	-2.1770*** (0.0687)

Notes. *, ** and *** denote significance at the 5%, 1% and 0.1% levels, respectively. Prais-Winsten regression models with panel-corrected standard errors adjusted for heteroskedasticity and panel-wide first-order autocorrelation.

Figure 1. Joint effect of skill and effort on performance (Staw 1980).

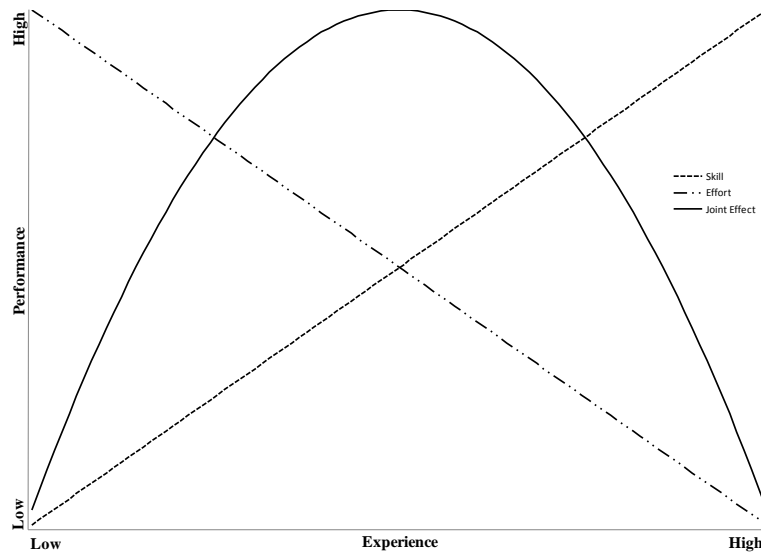


Figure 2. Process flow diagram for Shinsei loan process (parts of the process in white are included in the analyses).

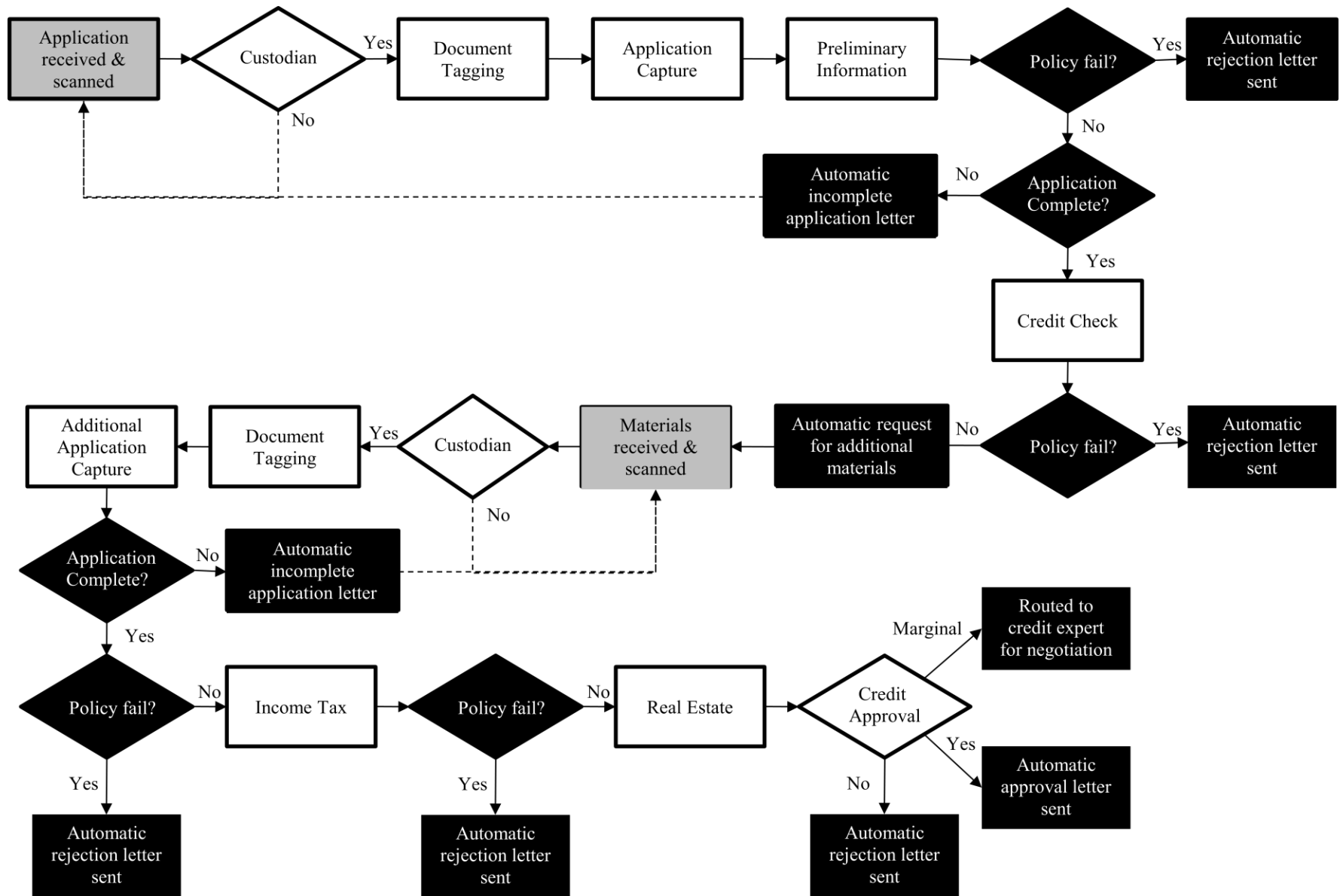
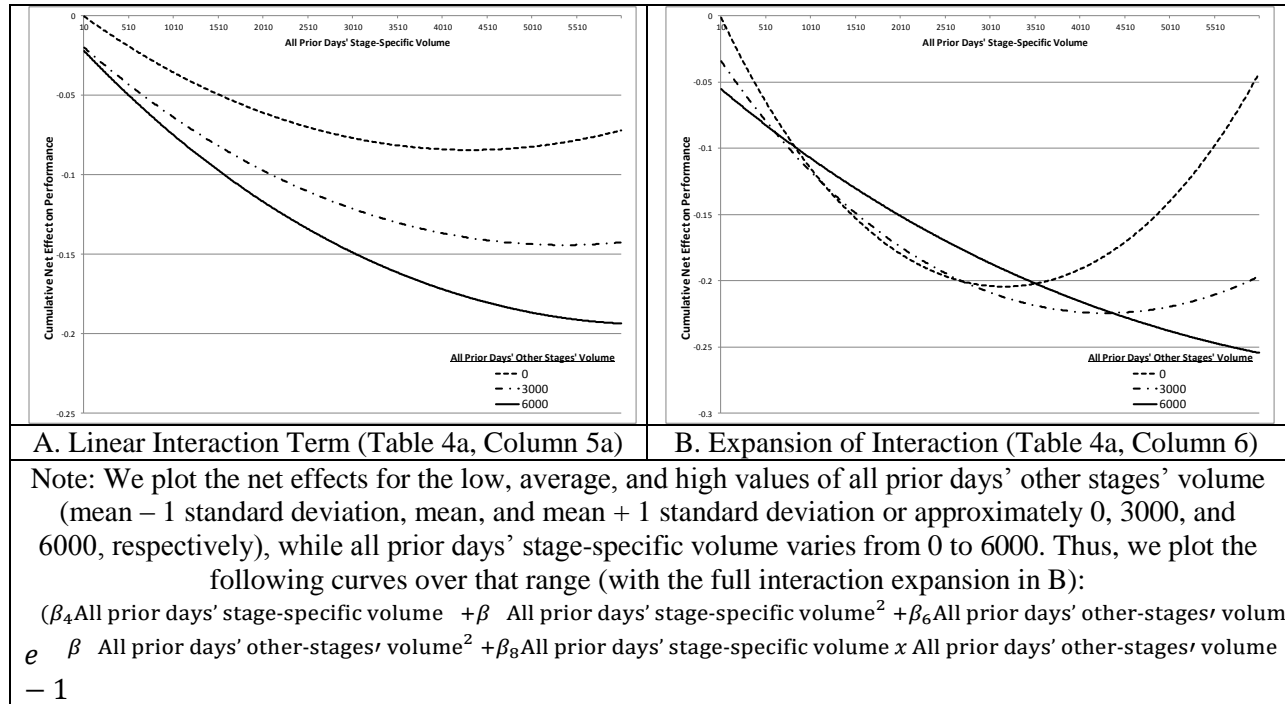


Figure 3. Examining the net effect of variety on performance.



6. References

- Allahverdi, A., J. N. D. Gupta and T. Aldowaisan (1999). "A review of scheduling research involving setup considerations." *Omega* **27**(2): 219-239.
- Allport, A., E. A. Styles and S. Hsieh (1994). Shifting intentional set. *Attention and Performance XV*. C. Umiltà and M. Moscovitch. Cambridge, MA, MIT Press: 421-452.
- Allport, A. and G. Wylie (2000). Task-switching, stimulus-response bindings and negative priming. *Control of Cognitive Processes: Attention and Performance XVIII*. S. Monsell and J. Driver. Cambridge, MA, MIT Press: 35-70.
- Anbalagan, P. and M. Vouk (2009). "'Days of the week' effect in predicting the time taken to fix defects." *Proceedings of the 2nd International Workshop on Defects in Large Software Systems*.
- Argote, L. (1999). *Organizational Learning: Creating, Retaining, and Transferring Knowledge*. Boston, Kluwer Academic.
- Argote, L., S. L. Beckman and D. Epple (1990). "The persistence and transfer of learning in industrial settings." *Management Sci.* **36**(2): 140-154.
- Argote, L. and E. Miron-Spektor (2011). "Organizational learning: From experience to knowledge." *Organ. Sci.*
- Argyris, C. and D. A. Schön (1978). *Organizational learning*. Reading, Mass., Addison-Wesley Pub. Co.
- Baddeley, A. (1992). "Working memory." *Science* **255**(5044): 556-559.
- Bahl, H. C., L. P. Ritzman and J. N. D. Gupta (1987). "Determining lot sizes and resource requirements: A review." *Operations Research* **35**(3): 329-345.
- Bailey, C. D. (1989). "Forgetting and the learning curve: A laboratory study." *Management Sci.* **35**(3): 340-352.
- Baum, J. A. C. and P. Ingram (1998). "Survival-enhancing learning in the Manhattan hotel industry, 1898-1980." *Management Sci.*: 996-1016.
- Beck, N. (2001). "Time-series cross-section data: What have we learned in the past few years?" *Annu. Rev. Polit. Sci.* **4**: 271-293.
- Bendoly, E., K. Donohue and K. L. Schultz (2006). "Behavior in operations management: Assessing recent findings and revisiting old assumptions." *J. of Operations Management* **24**(6): 737-752.
- Boh, W. F., S. A. Slaughter and J. A. Espinosa (2007). "Learning from experience in software development: A multilevel analysis." *Management Sci.* **53**(8): 1315-1331.
- Bohn, R. E. (2005). "From art to science in manufacturing: The evolution of technological knowledge." *Foundations and Trends in Technology, Information and Operations Management* **1**(2): 129-212.
- Bohn, R. E. and M. A. Lapré (2011). Accelerated learning by experimentation. *Learning Curves: Theory, Models, and Applications*. M. Y. Jaber. Boca Raton, FL, CRC Press, Taylor and Francis.
- Boudreau, J., W. Hopp, J. O. McClain and L. J. Thomas (2003). "On the interface between operations and human resources management." *Manufacturing Service Oper. Management* **5**(3): 179-202.
- Bryson, A. and J. Forth (2007). "Are there day of the week productivity effects?" *Center for Economic Performance, Manpower Human Resources Lab: #MHRLdp004*, London School of Economics.
- Cameron, A. C., J. B. Gelbach and D. L. Miller (Forthcoming). "Robust inference with multi-way clustering." *J. Bus. Econ. Stat.*
- Cellier, J.-M. and H. Eyrolle (1992). "Interference between switched tasks." *Ergonomics* **35**(1): 25 - 36.
- Clark, J. R. and R. S. Huckman (Forthcoming). "Broadening focus: Spillovers and the benefits of specialization in the hospital industry." *Management Sci.*
- Drucker, P. F. (1999). *Management Challenges for the 21st Century*. New York, HarperBusiness.
- Dutton, J. M. and A. Thomas (1984). "Treating progress functions as a managerial opportunity." *Acad. Management Rev.* **9**(2): 235-247.
- Ellis, H. C. (1965). *The transfer of learning*. New York, Macmillan.
- Fisher, C. D. (1993). "Boredom at work: A neglected concept." *Human Relations* **46**(3): 395-417.
- Fisher, M. L. and C. D. Ittner (1999). "The impact of product variety on automobile assembly operations: Empirical evidence and simulation analysis." *Management Sci.* **45**(6): 771-786.

- Fried, Y. and G. R. Ferris (1987). "The validity of the job characteristics model: A review and meta-analysis." *Pers. Psych.* **40**(2): 287-322.
- Gino, F., L. Argote, E. Miron-Spektor and G. Todorova (2010). "First, get your feet wet: The effects of learning from direct and indirect experience on team creativity." *Organ. Behav. Human Decision Processes* **111**(2): 102-115.
- Gino, F. and G. P. Pisano (2008). "Toward a theory of behavioral operations." *Manufacturing Service Oper. Management* **10**(4): 676-691.
- Gopher, D., L. Armony and Y. Greenshpan (2000). "Switching tasks and attention policies." *J of Experimental Psych: General* **129**(3): 308-339.
- Greenhalgh, L. and Z. Rosenblatt (1984). "Job insecurity: Toward conceptual clarity." *Acad. Management Rev.* **9**(3): 438-448.
- Hackman, J. R. (1969). "The nature of the task as a determiner of job behavior." *Pers. Psych.* **22**: 435-444.
- Hackman, J. R. and G. R. Oldham (1976). "Motivation through the design of work: Test of a theory." *Organ. Behav. Human Decision Processes* **16**(2): 250-279.
- Harrison, D. A. and K. J. Klein (2007). "What's the difference? Diversity constructs as separation, variety, or disparity in organizations." *Acad. Management Rev.* **32**(4): 1199-1228.
- Herzberg, F. (1968). "One more time: How do you motivate employees?" *Harvard Business Rev.* **46**(1): 53-62.
- Huckman, R. S. and A. MacCormack (2009). "BYD Company, Ltd." *HBS Case #606-139*.
- Huckman, R. S. and G. P. Pisano (2006). "The firm specificity of individual performance: Evidence from cardiac surgery." *Management Sci.* **52**(4): 473-488.
- Huckman, R. S. and B. R. Staats (2011). "Fluid tasks and fluid teams: The impact of diversity in experience and team familiarity on team performance." *Manufacturing Service Oper. Management* **13**(3): 310-328.
- Huckman, R. S., B. R. Staats and D. M. Upton (2009). "Team familiarity, role experience, and performance: Evidence from Indian software services." *Management Sci.* **55**(1): 85-100.
- Huckman, R. S. and D. Zinner (2008). "Does focus improve operational performance? Lessons from the management of clinical trials." *Strategic Management J.* **29**(2): 173-193.
- Humphrey, S. E., J. D. Nahrgang and F. P. Morgeson (2007). "Integrating motivational, social, and contextual work design features: A meta-analytic summary and theoretical extension of the work design literature." *J. Appl. Psych.* **92**(5): 1332-1356.
- Ichniowski, C. and K. Shaw (1999). "The effects of human resource management systems on economic performance: An international comparison of U.S. and Japanese plants." *Management Sci.* **45**(5): 704-721.
- Ingram, P. and T. Simons (2002). "The transfer of experience in groups of organizations: Implications for performance and competition." *Management Sci.*: 1517-1533.
- Kantor, P. B. and W. I. Zangwill (1991). "Theoretical foundation for a learning rate budget." *Management Sci.* **37**(3): 315-330.
- KC, D. and B. R. Staats (2011). "Learning from focal and varied experience: Evidence from Minimally Invasive Cardiac Surgery." *Working Paper*.
- KC, D. and C. Terwiesch (2009). "Impact of workload on service time and patient safety: An econometric analysis of hospital operations." *Management Sci.* **55**(9): 1486-1498.
- KC, D. and C. Terwiesch (Forthcoming). "The effects of focus on performance: Evidence from California hospitals." *Management Sci.*
- Langer, E. J. (1989). *Mindfulness*. Reading, MA, Addison-Wesley.
- Lapr  , M. A. (2011). Inside the learning curve: Opening the black box of the learning curve. *Learning Curves: Theory, Models, and Applications*. M. Y. Jaber. Boca Raton, FL, CRC Press, Taylor and Francis.
- Lapr  , M. A., A. S. Mukherjee and L. N. V. Wassenhove (2000). "Behind the learning curve: Linking learning activities to waste reduction." *Management Sci.* **46**(5): 597-611.

- Lapr , M. A. and I. M. Nembhard (2010). "Inside the organizational learning curve: Understanding the organizational learning process." *Foundations and Trends in Technology, Information and Operations Management* **4**(1): 1-103.
- Lapr , M. A. and N. Tsikriktsis (2006). "Organizational learning curves for customer dissatisfaction: Heterogeneity across airlines." *Management Sci.* **52**(3): 352-366.
- Levitt, B. and J. G. March (1988). "Organizational learning." *Annual Review of Sociology* **14**: 319-340.
- Levy, F. and R. J. Murnane (2004). The New Division of Labor: How Computers are Creating the Next Job Market. Princeton, N.J., Princeton University Press.
- Levy, F. K. (1965). "Adaptation in the production process." *Management Sci.* **11**(6): B136-B154.
- Lindbeck, A. and D. J. Snower (2000). "Multitask learning and the reorganization of work: From Tayloristic to holistic organization." *J. Labor Econ.* **18**(3): 353-376.
- MacDuffie, J. P., K. Sethuraman and M. L. Fisher (1996). "Product variety and manufacturing performance: Evidence from the international automotive assembly plant study." *Management Sci.* **42**(3): 350-369.
- Monsell, S. (2003). "Task switching." *Trends in cognitive sciences* **7**(3): 134-140.
- Mukherjee, A., W. Mitchell and F. B. Talbot (2000). "The impact of new manufacturing requirements on production line productivity and quality at a focused factory." *J. of Operations Management* **18**: 139-168.
- Narayanan, S., S. Balasubramanian and J. M. Swaminathan (2009). "A matter of balance: Specialization, task variety, and individual learning in a software maintenance environment." *Management Sci.* **55**(11): 1861-1876.
- Nembhard, D. A. and N. Osothsilp (2001). "An empirical comparison of forgetting models." *IEEE Transactions on Engineering Management* **48**(3): 283-291.
- Newell, A. and P. Rosenbloom (1981). Mechanisms of skill acquisition & the power law of practice. Cognitive Skills & Their Acquisition. J. Anderson. Hillsdale, NJ, Erlbaum: 1-55.
- Ortega, J. (2001). "Job rotation as a learning mechanism." *Management Sci.* **47**(10): 1361-1370.
- Piaget, J. (1963). The Psychology of Intelligence. New York, Routledge.
- Pines, A. and C. Maslach (1978). "Characteristics of staff burnout in mental health settings." *Psychiatric Services* **29**(4): 233.
- Quinn, R. W. (2005). "Flow in knowledge work: High performance experience in the design of national security technology." *Administrative Sci. Q.* **50**(4): 610-641.
- Reagans, R., L. Argote and D. Brooks (2005). "Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together." *Management Sci.* **51**(6): 869-881.
- Roethlisberger, F. J. and W. J. Dickson (1934). Management and the Worker. Boston, Harvard University.
- Rogers, R. D. and S. Monsell (1995). "Costs of a predictable switch between simple cognitive tasks." *J of Experimental Psych.* **124**: 207-231.
- Roy, D. F. (1959). "'Banana time': Job satisfaction and informal interaction." *Human Organization* **18**(4): 158-168.
- Rubinstein, J. S., D. E. Meyer and J. E. Evans (2001). "Executive control of cognitive processes in task switching." *J of Experimental Psych: Human Perception and Performance* **27**(4): 763.
- Sarason, I. G. and G. R. Pierce (1996). Cognitive Interference, Lawrence Erlbaum Associates.
- Schendel, J. D. and J. D. Hagman (1982). "On sustaining procedural skills over a prolonged retention interval." *J. Appl. Psych.* **67**(5): 605-610.
- Schilling, M. A., P. Vidal, R. E. Ployhart and A. Marangoni (2003). "Learning by doing something else: Variation, relatedness, and the learning curve." *Management Sci.* **49**(1): 39-56.
- Schmidt, R. A. (1975). "A schema theory of discrete motor skill learning." *Psych. Rev.* **82**(4): 225-260.
- Schultz, K. L., D. C. Juran, J. W. Boudreau, J. O. McClain and L. J. Thomas (1998). "Modeling and worker motivation in JIT production systems." *Management Sci.* **44**(12): 1595-1607.

- Schultz, K. L., J. O. McClain and L. J. Thomas (2003). "Overcoming the dark side of worker flexibility." *J. of Operations Management* **21**(1): 81-92.
- Schultz, K. L., T. Schoenherr and D. Nembhard (2010). "An example and a proposal concerning the correlation of worker processing times in parallel tasks." *Management Sci.* **56**(1): 176-191.
- Scott, W. E. (1966). "Activation theory and task design." *Organ. Behav. Human Decision Processes* **1**(1): 3-30.
- Shingo, S. (1989). A Study of the Toyota Production System from an Industrial Engineering Viewpoint. Cambridge, MA, Productivity Press.
- Simons, J. V. and G. R. Russell (2002). "A case study of batching in a mass service operation." *J. of Operations Management* **20**(5): 577-592.
- Simonsohn, U. (Forthcoming). "Weather to go to college." *The Economic Journal*.
- Skinner, W. (1974). "The focused factory." *Harvard Business Rev.* **52**(3): 113.
- Skinner, W. (1985). The taming of lions: How manufacturing leadership evolved. The Uneasy Alliance: Managing the Productivity-Technology Dilemma. K. B. Clark, R. H. Hayes and C. Lorenz. Boston, HBS Press.
- Spear, S. and H. K. Bowen (1999). "Decoding the DNA of the Toyota Production System." *Harvard Business Rev.* **77**(5): 97-106.
- Staw, B. M. (1980). "The consequences of turnover." *J. of Occupational Behav.* **1**(4): 253-273.
- Stross, R. (2010). When the Assembly Line Moves Online. New York Times. New York.
- Taylor, F. W. (1911). The Principles of Scientific Management. New York, Harper & Brothers.
- Tsikriktsis, N. (2007). "The effect of operational performance and focus on profitability: A longitudinal study of the U.S. airline industry." *Manufacturing Service Oper. Management* **9**(4): 506-517.
- Tucker, A. L., I. M. Nembhard and A. C. Edmondson (2007). "Implementing new practices: An empirical study of organizational learning in hospital intensive care units." *Management Sci.* **53**(6): 894-907.
- Tzur, M. (1996). "Learning in setups: Analysis, minimal forecast horizons, and algorithms." *Management Sci.* **42**(12): 1732-1743.
- Upton, D. M. and J. D. Margolis (1992). "McDonald's Corp." *HBS Case #693-028*.
- Upton, D. M. and B. R. Staats (2008). "Radically Simple IT." *Harvard Business Rev.* **86**(3): 118-124.
- Vroom, V. H. (1964). Work and Motivation. New York, Wiley.
- Warr, P. (2007). Work, Happiness, and Unhappiness. Mahwah, NJ, Erlbaum.
- Waszak, F., B. Hommel and A. Allport (2003). "Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs." *Cognitive Psych.* **46**(4): 361-413.
- Wiersma, E. (2007). "Conditions that shape the learning curve: Factors that increase the ability and opportunity to learn." *Management Sci.* **53**(12): 1903-1915.
- Wylie, G. and A. Allport (2000). "Task switching and the measurement of switch costs." *Psych. Research* **63**: 212-233.
- Yelle, L. E. (1979). "The learning curve: Historical review and comprehensive story." *Decision Sci.* **10**(2): 302-328.